

DOCUMENT RESUME

ED 236 175

TM 830 676

AUTHOR Green, Donald Ross; Yen, Wendy M.
TITLE Number Correct versus Pattern Scoring: Results for Ethnic Groups.
PUB DATE Apr 83
NOTE 31p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Montreal, Quebec, April 12-14, 1983).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Basic Skills; *Black Students; Elementary Education; Guessing (Tests); *Scoring; Standardized Tests; Test Bias; *Testing Problems; Testing Programs
IDENTIFIERS Comprehensive Tests of Basic Skills; *Number Right Scoring; *Pattern Scoring

ABSTRACT

The Comprehensive Tests of Basic Skills, Form U, is scored in two ways: number-correct and pattern. The latter makes use of the information about which particular items are answered correctly, giving more weight to the more discriminating items and making allowances for guessing. Critics have suggested that black students are penalized by pattern scoring on the ground that they tend to guess more than others. For this study results for two tests at each of four different levels (and grades) of the battery were used. The two tests were Reading Comprehension and Mathematics Concepts and Applications. The data used were drawn from both the 1980 Fall and 1981 Spring standardizations of levels D, F, G, and H of the battery given at grades 2, 4, 6, and 8 respectively. A comparison of the mean differences between the two types of scores for blacks (n=4760) and others (n=25,280) did not support the contention that blacks are penalized by pattern scoring.
(Author/PN)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED236175

Number Correct Versus Pattern Scoring:
Results for Ethnic Groups

Donald Ross Green
CTB/McGraw-Hill

Wendy M. Yen
CTB/McGraw-Hill

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

D. R. Green

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

Presented at the meeting of the National Council
on Measurement in Education, Montreal, April 1983

Abstract

The Comprehensive Tests of Basic Skills, Form U is scored in two ways: number-correct and pattern. The latter makes use of the information about which particular items are answered correctly, giving more weight to the more discriminating items and making allowances for guessing. Critics have suggested that black students are penalized by pattern scoring on the ground that they tend to guess more than others. A comparison of the mean differences between the two types of scores for blacks and others did not support this contention.

When a test has been calibrated using the three parameter model it is possible to score the test and obtain a trait estimate based on the pattern of item scores (Lord, 1980). This item pattern scoring produces somewhat more reliable scores than one can obtain using only the number correct information (Yen, 1983). Trait estimates derived from pattern scoring are closely related to those derived from number correct scoring since the number of items correct is the starting point for estimating pattern trait values. However, pattern scoring further takes into consideration which items are answered correctly, giving more weight to the more discriminating items and making allowances for the probability of guessing. It is this latter feature that gives rise to the present investigation.

When CTBS/U was published in the Fall of 1981 and the availability of pattern scoring services was announced by CTB, some other test publishers and a number of school people raised questions about the effect of this scoring procedure on the scores of groups that traditionally have large numbers falling within or near the so-called "chance" score range. The fear has been expressed that somehow these groups would be unduly penalized by pattern scoring. Specifically some said that blacks, Hispanics, and minority children generally would score unusually low. The unstated assumption is that members of these groups guess more often than those of other groups although there does not appear

to be any evidence to support this assumption. That is, the critics seem to be saying that given two groups of low scoring students whose performance appears equal on a number correct basis, the minority students will be penalized by a scoring system which considers certain score patterns to indicate less ability than other patterns. Of course this amounts to saying that number correct scoring probably overestimates the performance of minority students in comparison with majority students of equal achievement. Again there does not seem to be any evidence to support this notion.

It is true that, on the average, school children in minority groups tend to obtain lower scores on standardized tests than other children. It is also true that schools often give "on-level" tests to all students with the consequence that low performers tend to be given tests on which their true scores are at or near the "chance" level on the tests. It is commonly assumed that this means that low scoring children do nothing but guess (e.g., Fry, 1971), all of which apparently leads some people to the conclusion that minority and disadvantaged school children tend to guess on these tests more often than do other children with equal levels of achievement.

Of course not all schools give only on-level tests, not all minority children are low scoring, and furthermore it does not appear to be true that most children whose scores fall in the so-called chance range of a test respond randomly (McRae & Green, 1971). On the other hand low scoring students tend to guess more often than high scoring students on the average and thus guessing

does play a larger role in the observed scores of the low scorers (Doscher & Bruno, 1981). While none of this suggests differential guessing among ethnic groups given equal true score levels, such a phenomenon is possible. This study will explore that possibility by looking for differences between number correct and pattern scores within two ethnic groups.

Method

The Comprehensive Tests of Basic Skills, Form U (CTBS/U; CTB/McGraw-Hill, 1981) were standardized in the Fall of 1980 and the Spring of 1981. For this study results for two tests at each of four different levels (and grades) of the battery were used. The two tests are Reading Comprehension and Mathematics Concepts and Applications. The data used were drawn from both the Fall and Spring standardizations of levels D, F, G, and H of the battery given at grades 2, 4, 6, and 8 respectively.

All students participating in the standardization were asked to identify themselves as Black, Hispanic or Other. The data used were the scores obtained by the Black and Other students who reported this information. The number of cases in these groups in each test at each level are shown in Table 1.

Each test was scored in two ways: first a pattern score ability estimate (theta) for each student was obtained using LOGIST (Wood, Wingersky, & Lord, 1976); this is a maximum likelihood estimate. A second maximum likelihood estimate of student

TABLE 1
Numbers of Students

Grade/Level		Blacks	Others
2	D	1,014	8,190
4	F	859	4,887
6	G	1,137	6,445
8	H	1,750	5,758

ability was determined based solely on the number of items correct using the procedure developed by Yen. The scoring procedures used are those in use at CTB for these two types of scores. Thus omits and not reaches were treated as wrong answers. The procedures for establishing floors and ceilings described by Yen in her paper were also used (Yen, 1983).

Next an average of these two scores was obtained for each student and this average was used to assign the students in each ethnic group to one of four score groups running from low to high. The lowest score group consisted of those students whose average score was at or below "chance" level as determined from the raw score to theta table obtained in the number correct scoring procedure. The highest score level was set so that at least 100 cases would be included in the smaller (Black) group. The difference was split for the two remaining score groups.

The mean difference between the two thetas, pattern and number correct, was determined separately for each group in each score range. The overall means, standard deviations, and correlations for each test at each level for the two sorts of scores were also determined for each group.

Results

The correlations are shown in Table 2. Clearly they are all very high and they corroborate for the two groups what Yen (1983) has reported for all students taken together. There does seem to be slight tendency for the correlations for the Black group to be

TABLE 2
Correlation Coefficients

Grade/Level		Test	Blacks	Others
2	D	Reading Comprehension	.96	.98
		Math Concepts & Applications	.97	.98
4	F	Reading Comprehension	.97	.98
		Math Concepts & Applications	.93	.96
6	G	Reading Comprehension	.98	.99
		Math Concepts & Applications	.99	.99
8	H	Reading Comprehension	.98	.98
		Math Concepts & Applications	.96	.99

lower than those for the Other group. This is presumably because of the somewhat larger proportions of the blacks scoring in the interval near and below chance where the relationship between the two scores is attenuated.

The mean thetas for the two scores and the differences between these two means are shown for the two ethnic groups by score group for each test at each grade (test level) in Tables 3 - 10. The variances for these sets of trait estimates are not all the same so the mean differences are not quite comparable. Therefore pooled estimates of the standard deviations for the two sets of thetas were obtained by finding the square root of their mean variance. These are shown in Table 11. The mean differences shown Tables 3 - 10 were divided by these standard deviations to obtain the standardized differences between the means listed in Table 12.

These standardized differences tend to be quite small. Most of the larger ones come from the smallest groups. For the Black group pattern scoring usually produces very slightly higher mean scores (negative differences), particularly at the lower score levels. For the Other group there is no overall difference between the scoring methods, although in the lowest score level most of the nonzero differences are negative.

Discussion

The present data offer no support for the possibility that black students would be penalized by pattern scoring. If anything the data show that pattern scoring is slightly more favorable to the minority group, but this trend is not strong enough to say

TABLE 3
Mean Number Correct and Mean Pattern Scores
by Interval for Grade 2 Reading Comprehension

INTERVAL	Blacks					Others				
	N	NR0	PATT0	DIFF	SD _{DIFF}	N	NR0	PATT0	DIFF	SD _{DIFF}
-4.00 - -1.88	291	-2.64	-2.68	0.05	0.61	1,292	-2.66	-2.68	0.03	0.58
-1.87 - -0.75	223	-1.20	-1.19	-0.01	0.25	970	-1.21	-1.19	-0.02	0.27
-0.74 - + .40	331	-0.18	-0.17	-0.01	0.19	2,672	-0.07	-0.09	0.02	0.11
+0.41 - +4.00	169	0.99	0.93	0.06	0.17	3,256	1.12	1.07	0.04	0.20
TOTALS	1,014	-0.92	-0.93	0.02	0.37	8,190	-0.14	-0.17	0.03	0.29

TABLE 4

Mean Number Correct and Mean Pattern Scores
by Interval for Grade 2 Math Concepts & Applications

INTERVAL	Blacks					Others				
	N	NRØ	PATTØ	DIFF	SD _{DIFF}	N	NRØ	PATTØ	DIFF	SD _{DIFF}
-4.00 - -3.39	79	-3.72	-3.67	-0.04	0.20	259	-3.73	-3.71	-0.02	0.15
-3.38 - -1.50	212	-2.21	-2.11	-0.10	0.58	502	-2.18	-2.04	-0.14	0.52
-1.49 - +0.40	587	-0.55	-0.51	-0.04	0.18	4,207	-0.34	-0.35	0.01	0.18
+0.41 - +4.00	136	1.09	1.10	-0.01	0.31	3,222	1.32	1.26	0.05	0.34
TOTALS	1,014	-0.92	-0.88	-0.05	0.32	8,190	0.09	0.08	0.01	0.29

TABLE 5
Mean Number Correct and Mean Pattern Scores
by Interval for Grade 4 Reading Comprehension

INTERVAL	Blacks					Others				
	N	NRØ	PATTØ	DIFF	SD _{DIFF}	N	NRØ	PATTØ	DIFF	SD _{DIFF}
-4.00 - -2.43	351	-2.74	-2.72	-0.02	0.13	524	-2.72	-2.70	-0.02	0.18
-2.42 - -1.26	153	-1.86	-1.90	0.04	0.67	494	-1.75	-1.89	0.14	0.65
-1.25 - -0.10	204	-0.67	-0.68	0.02	0.12	1,224	-0.59	-0.60	0.01	0.11
-0.09 - +4.00	151	0.40	0.42	-0.02	0.20	2,645	0.74	0.73	0.01	0.11
TOTALS	859	-1.54	-1.54	0.00	0.32	4,887	-0.22	-0.24	0.02	0.24

TABLE 6

Mean Number Correct and Mean Pattern Scores
by Interval for Grade 4 Math Concepts & Applications

INTERVAL	Blacks					Others				
	N	NR0	PATT0	DIFF	SD _{DIFF}	N	NR0	PATT0	DIFF	SD _{DIFF}
-4.00 - -2.98	71	-3.31	-3.82	-0.02	-0.22	267	-3.31	-3.30	-0.01	0.19
-2.97 - -1.42	163	-2.16	-1.95	-0.21	1.00	488	-2.14	-2.15	0.01	1.00
-1.41 - +0.15	435	-0.63	-0.61	-0.02	0.19	1889	-0.48	-0.47	-0.02	0.19
+0.16 - +4.00	190	0.65	0.64	0.01	0.12	2243	0.88	0.87	0.01	0.13
TOTALS	859	-0.86	-0.81	-0.05	0.47	4887	-0.18	-0.18	0.00	0.35

TABLE 7
Mean Number Correct and Mean Pattern Scores
by Interval for Grade 6 Reading Comprehension

INTERVAL	Blacks					Others				
	N	NR0	PATT0	DIFF	SD _{DIFF}	N	NR0	PATT0	DIFF	SD _{DIFF}
-4.00 - -2.97	384	-3.10	-3.10	0.00	0.03	561	-3.09	-3.09	0.00	0.05
-2.96 - -1.41	196	-2.09	-2.07	-0.02	0.64	561	-2.14	-2.09	-0.05	0.66
-1.40 - +0.15	408	-0.59	-0.57	-0.02	0.13	2167	-0.46	-0.47	0.01	0.14
+0.16 - +4.00	149	0.63	0.64	-0.01	0.15	3156	0.87	0.87	0.00	0.09
TOTALS	1137	-1.54	-1.52	-0.01	0.28	6445	-0.19	-0.19	0.00	0.22

TABLE 8

Mean Number Correct and Mean Pattern Scores
by Interval for Grade 6 Math Concepts & Applications

INTERVAL	Blacks					Others				
	N	NR0	PATT0	DIFF	SD _{DIFF}	N	NR0	PATT0	DIFF	SD _{DIFF}
-4.00 - -3.01	297	-3.70	-3.67	-0.03	0.33	473	-3.69	-3.67	-0.03	0.37
-3.00 - -1.43	157	-1.97	-2.05	0.08	0.61	507	-1.95	-2.02	0.08	0.62
-1.42 - +0.15	520	-0.60	-0.58	-0.03	0.13	2406	-0.49	-0.48	-0.01	0.13
+0.16 - +4.00	163	0.61	0.64	-0.03	0.21	3059	0.91	0.90	0.01	0.13
TOTALS	1137	-1.43	-1.42	-0.01	0.31	6445	-0.18	-0.18	0.01	0.24

TABLE 9
Mean Number Correct and Mean Pattern Scores
by Interval for Grade 8 Reading Comprehension

INTERVAL	Blacks					Others				
	N	NR0	PATT0	DIFF	SD _{DIFF}	N	NR0	PATT0	DIFF	SD _{DIFF}
-4.00 - -3.00	735	-3.33	-3.31	-0.02	0.09	636	-3.33	-3.31	-0.02	0.09
-2.99 - -1.54	200	-2.28	-2.04	-0.23	0.74	386	-2.12	-2.21	0.09	0.77
-1.53 - +0.40	677	-0.55	-0.54	0.00	0.17	2566	-0.27	-0.27	0.00	0.15
+0.41 - +4.00	138	0.91	0.87	0.04	0.16	2170	1.02	1.02	0.00	0.22
TOTALS	1750	-1.80	-1.77	-0.03	0.29	5758	-0.24	-0.25	0.01	0.27

TABLE 10

Mean Number Correct and Mean Pattern Scores
by Interval for Grade 8 Math Concepts & Applications

INTERVAL	Blacks					Others				
	N	NRØ	PATTØ	DIFF	SD _{DIFF}	N	NRØ	PATTØ	DIFF	SD _{DIFF}
-4.00 - -2.48	229	-3.35	-3.21	-0.13	0.64	595	-3.41	-3.43	0.01	0.45
-2.47 - -0.91	510	-1.52	-1.45	-0.07	0.39	662	-1.43	-1.46	0.02	0.40
-0.90 - +0.40	759	-0.29	-0.29	0.00	0.12	2092	-0.15	-0.14	-0.01	0.12
+0.41 - +4.00	252	0.88	0.96	-0.08	0.27	2409	1.03	1.03	0.00	0.11
TOTALS	1750	-0.88	-0.83	-0.05	0.34	5758	-0.14	-0.14	0.00	0.22

TABLE 11

Pooled Standard Deviations

Grade/Level		Test	Blacks	Others
2	D	Reading Comprehension	1.35	1.39
		Math Concepts & Applications	1.37	1.36
4	F	Reading Comprehension	1.25	1.30
		Math Concepts & Applications	1.58	1.32
6	G	Reading Comprehension	1.40	1.36
		Math Concepts & Applications	1.23	1.42
8	H	Reading Comprehension	1.53	1.48
		Math Concepts & Applications	1.29	1.44

TABLE 12
Standardized Mean Differences Between
Number Correct Thetas and Pattern Thetas*

			Interval 1		Interval 2		Interval 3		Interval 4		Total	
Grade/Level		Test	Blacks	Others	Blacks	Others	Blacks	Others	Blacks	Others	Blacks	Others
2	D	Rdg Comp	0.04	0.02	0.00	-0.01	0.00	0.01	0.04	0.03	0.01	0.02
		Math C&A	-0.03	-0.01	-0.07	-0.10	-0.03	0.00	0.00	0.04	-0.03	0.01
4	F	Rdg Comp	-0.02	-0.02	0.03	0.11	0.02	0.00	-0.02	0.00	0.00	0.02
		Math C&A	-0.01	0.00	-0.13	0.00	-0.01	-0.02	0.00	0.00	-0.03	0.00
6	G	Rdg Comp	0.00	0.00	-0.01	-0.04	-0.01	0.00	0.00	0.00	0.00	0.00
		Math C&A	-0.02	-0.02	0.07	0.06	-0.02	0.00	-0.02	0.00	-0.01	0.00
8	H	Rdg Comp	-0.01	-0.01	-0.15	0.06	0.00	0.00	0.03	0.00	-0.02	0.00
		Math C&A	-0.10	0.00	-0.05	0.01	0.00	0.00	-0.06	0.00	-0.04	0.00

*A positive difference indicates that the mean of the number correct thetas is higher than the mean of the pattern thetas.

that pattern scoring systematically favors blacks. Pattern scoring may be slightly more favorable to low scoring students generally although no logical reason for this is apparent.

In general the results tend to confirm for the two groups separately what Yen has found for all groups considered together. The correlations between the two types of scores are very high. In fact the correlations are so high that, coupled with the similarity of the means, they should put to rest any fears that the choice of one or the other of the two scoring procedures might introduce some systematic distortion in the results of a testing program.

So far then, the only differences that have been found between pattern scoring and number correct scoring is that the former yields somewhat smaller standard errors of measurement, particularly for low scoring students.

References

- Comprehensive Tests of Basic Skills, Form U. Monterey, CA: CTB/McGraw-Hill, 1981.
- Doscher, M. and Bruno, J.G. Simulation of inner-city standardized testing behavior: implications for instructional evaluation. American Educational Research Journal, 1981, 18, 475-488.
- Fry, E. The orangoutang score. The Reading Teacher, 1971, 24, 360-362.
- Lord, F.M. Applications of item response theory to practical testing problems. New York: Lawrence Erlbaum Associates, Inc., 1980.
- McRae, D.J. and Green, D.R. Comment on "The orangoutang score". The Reading Teacher, 1971, 24, 364.
- Wood, R.L., Wingersky, M.S., and Lord, F.M. LOGIST: A computer program for estimating examinee ability and item characteristic curve parameters (RM-76-6). Princeton, NJ: Educational Testing Service, 1976.
- Yen, W.M. Obtaining maximum likelihood trait estimates from number-correct scores for the three-parameter logistic model. Paper presented at the American Educational Research Association meetings, Montreal, April 1983.